PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12P 19/34, C12N 15/00 C12Q 1/68

A1

(11) International Publication Number: WO 90/01064

(43) International Publication Date: 8 February 1990 (08.02.90)

(21) International Application Number: PCT/US89/03099

(22) International Filing Date: 18 July 1989 (18.07.89)

(30) Priority data: 225,037 26 July 1988 (26.07.88) US

(71) Applicant: GENELABS INCORPORATED [US/US]; 505 Penobscot Drive, Redwood City, CA 94063 (US).

(72) Inventors: LARRICK, James ; Star Route Box 48, Woodside, CA 94062 (US). FRY, Kirk ; 2604 Ross Road, Palo Alto, CA 94303 (US). TAM, Albert ; 1871 10th Avenue, San Francisco, CA 94122 (US).

(74) Agent: DEHLINGER, Peter, J.; 350 Cambridge Avenue, Suite 100, Palo Alto, CA 94306 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FR (European patent), GB (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), SE (European patent).

Published

With international search report.
With amended claims.

(54) Title: SEQUENCE-SPECIFIC AMPLIFICATION TECHNIQUES

(57) Abstract

A method of selectively amplifying nucleic acid fragments having a known sequence. Sense-strand fragments containing the known sequence are primed with a sequence-specific primer having a sequence which is homologous to the known sequence, and replicated in the presence of DNA polymerase to form anti-sense strands having the primer sequence at their 5' ends. The anti-sense strands are treated with terminal deoxynucleotide transferase and a selected deoxynucleoside triphosphate, to add a homopolymeric sequence to the 3' strand ends. The resulting anti-sense strand fragments are mixed with a homopolymer primer which is homologous to said homopolymeric fragment sequence, a common-sequence primer which is homologous to a region of the specific-sequence primer, DNA polymerase and all four deoxynucleoside triphosphates. The newly synthesized fragments are amplified with repeated cycles of denaturing annealing, and primed polymerization. Also disclosed is a method for identifying and analyzing restriction fragment length polymorphisms, which employs the fragment amplification method.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Fasso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

SEQUENCE-SPECIFIC AMPLIFICATION TECHNIQUES

1. Field of the Invention

The present invention relates to methods of selectively amplifying sequence specific RNA and single- or double-stranded DNA fragments.

2. References

Cathala, G., et al, DNA (1983) 2(4):329.

Holbrook, N.J., et al, Proc Natl Acad Sci (USA) (1984) 81:1634.

Jackson, D.A., et al, Proc Natl Acad Sci (1972) 69:2904.

Lobban, P.E., et al, J Mol Biol (1973) 78:453.

Maniatis, T., et al, Molecular Cloning: Laboratory Manual, Cold Spring Harbor Laboratory (1982), p. 280.

Nelson, T. et al, Methods in Enzymology, Academic Press, NY (1979) 68:41-50.

Roychoudhury, R.E., et al, Nucleic Acids Res (1976) 3:101.

Thompson, J., et al, Anal Biochem (1987) 303:334.

3. Background of the Invention

Several techniques have been proposed for identifying or isolating specific-sequence RNA or DNA fragments from a fragment mixture. These techniques may be used, for example, to identify the presence of disease-

related nucleic acid species in a cellular or tissue sample, or to isolate a specific gene or RNA transcript in mixture of cellular or viral RNA or DNA fragments.

One method for selecting a specific sequence from a mixture involves reacting the RNA or DNA species with a complementary-sequence probe which may be bound to a solid support, and/or which carries a label for quantitating probe binding to the nucleic acid fragments. Another method is designed for isolating sequences which are unique to one of two sources of fragment material. In this method, fragments from the two sources are mixed, denatured, reannealed, then separated on the basis of cross-hybridization (common sequences) and self-hybridization (unique sequences). Both of these methods are limited by the amount of fragment material obtainable from the fragment sources, or involve intermediate cloning steps to increase the amount of fragment material.

More recently, a method for amplifying duplex DNA fragments by repeated strand-replication has been described (U.S. Patents Nos. 4,683,194 and 4,683,202). This method, generally referred to as a polymerase chain reaction (PCR) method, is designed to selectively amplify fragments whose sequences are complementary to two different sequence-specific primers. One limitation of the previously described PCR method is the need for two sequence-specific primers. Also, the polymerization chain reaction requires a relatively high concentration of both sequence-specific primers. Another limitation of the earlier-described method is that only the region between the two probes is amplified, and therefore identification of full-length fragments, or fragment lengths which are determined by restriction sites outside of the probe regions is not possible.

4. Summary of the Invention

It is therefore one object of the invention to provide a method of identifying and/or isolating sequence-specific fragments which overcomes the problems and limitations associated with prior art methods discussed above.

It is a more specific object of the invention to provide such a method which allows fragment selection using a single sequence-specific probe, and which allows amplification without sequence-specific probes.

A related object of the invention is to provide such a method which allows sequence specific fragment amplification with a single homopolymeric probe.

Still another object of the invention is to provide a method for amplifying the region of a sequence-specific RNA or DNA fragment from a known 3' sequence region to the 5' end of the fragment, such as 5'-end restriction sites. In particular, the method may be used to amplify and identify sequence-specific restriction fragment length polymorphisms.

Another object of the invention is to provide such a method which can produce amplified fragments with selected end restriction sites for either directional or random-direction cloning.

The invention provides a method of selectively amplifying nucleic acid fragments having a region of known sequence. The method includes mixing sense-strand fragments containing the known sequence with a sequence-specific primer having a 3' sequence which is homologous to the known sequence, and reacting the primed fragments under strand synthesis conditions. The resulting anti-sense strands, which have a common primer sequence at their 5' ends, are treated with terminal deoxynucleotide transferase (TdT) and a selected deoxynucleoside triphosphate, to add a homopolymeric sequence to the 3'

ends of the anti-sense strand fragments. The fragments are mixed with a homopolymer primer which is homologous to a 3'-end homopolymeric sequence in the anti-sense strands, a known- or common-sequence primer which is homologous to the complement of the 5'-end common primer sequence, DNA polymerase, and all four deoxynucleoside triphosphates. The sequence-specific fragments are amplified by repeated cycles of primer annealing, polymerization, and denaturation.

The method may be used for amplifying duplex fragments, such as genomic fragments, having a region of known sequence. Here the duplex fragments are preferably first blocked at their 3' ends, prior to initial anti-sense strand formation, to prevent homopolymer tailing of the original DNA strands. In one embodiment, the specific-sequence primer contains a 5' sequence which is complementary to the homopolymeric sequence added at the 3' end of the anti-sense strand, and the homopolymer primer also serves as the common-sequence primer, i.e., only a single primer is required for amplification.

The method is also useful for identifying restriction fragment length polymorphisms in duplex DNA fragments, where the fragments of interest have a region with the same known sequence, and specific restriction endonuclease sites which are spaced at different distances from the known sequence region. After amplifying the different restriction length fragments, fragment length differences are observed by size fractionating the fragments, such as by gel electrophoresis.

The method may also be used for placing the same or different selected restriction endonuclease sites at the ends of the amplified fragments, for non-directional or directional fragment cloning. The selected restriction site sequences are included at the 5' ends of the homopolymer and common-sequence primers.

In another general aspect, the method is designed for selectively amplifying anti-sense strand fragments derived from single-strand RNA, typically cDNA strands derived from messenger RNA (mRNA). Here the sequence-specific primer is mixed with the RNA species, to produce an anti-sense strand cDNA which has a common primer sequence at its 5′ end. In one embodiment, the 5′-end sequence of the specific-sequence primer is complementary to the homopolymeric sequence which is added at the 3′ ends of the anti-sense strands, and the homopolymer primer also serves as the defined-sequence primer, i.e., amplification requires only a single primer. The amplified fragments may be equipped with selected restriction site ends, as above.

These and other objects and features of the invention will become more fully appreciated when the following detailed description of the invention is read in conjunction with the accompanying drawings.

Brief Description of the Drawings

Figure 1 is a flow diagram of a method of selectively amplifying an RNA species having a region R of known sequence, according to one embodiment of the method of the invention;

Figure 2 shows how the amplification method of Figure 1 can be adapted for use in directional cloning of the amplified fragments;

Figure 3 is a flow diagram of a method of selectively amplifying a duplex DNA fragment having a region R of known sequence, according to another embodiment of the method of the invention; and

Figure 4 illustrates the use of the fragment amplification method for detecting restriction length fragment polymorphisms.

Detailed Description of the Invention

I. Selective Amplification: mRNA Species

The method of the invention is useful for selectively amplifying DNA fragments which are derived from single-strand RNA species, typically messenger RNA (mRNA) species, having a region of known sequence. Methods for isolating mRNA species from tissue, cellular or body-fluid samples are well known. One method involves formation of a vanadyl-RNA complex, extraction of protein with chloroform/phenol, and precipitation with cold ethanol. In a second method, the RNA is extracted from a guanidium isothiocyanate mixture with phenol, followed by a chloroform:isoamyl alcohol extraction, and precipitation of RNA from the aqueous phase with cold ethanol. The reader is referred to Maniatis, pp. 188-198, and references cited therein for details. One preferred RNA isolation method is described in Cathala.

Most eukaryotic mRNAs are characterized by a 3' terminal poly A sequence which allows isolation by affinity chromatography, using oligo dT bound to a solid support. In addition, or alternatively, total isolated RNA can be further fractionated by density gradient centrifugation, or agarose gel electrophoresis, to obtain a desired size fraction of RNA species.

Figure 1 shows a poly A-containing mRNA (poly A RNA) species having a region R of known sequence. This species is typically present as one of a large number of other RNA species with different sequences. As a first step in the method, the RNA species are mixed with a sequence-specific primer having a region which is homologous to the known sequence in region R. The primer may also include a region which has an unrelated sequence, preferably a homopolymeric sequence which is 5' with respect to the sequence-specific primer region, such as a

poly dC sequence, as indicated in the figure. The region of the primer which is homologous to the known sequence is indicated by a sawtooth line in the figure, and typically includes at least about 20, and up to several hundred or more specific bases. Where the primer includes an additional homopolymer sequence, such preferably includes at least about 14-20 homopolymer bases.

The known-sequence RNA, which is primed by the added primer, is copied under standard conditions by reverse transcriptase in the presence of all four deoxynucleoside triphosphates (dNTPs). As seen in Figure 1, this reaction produces an anti-sense cDNA strand which extends from the known-sequence region of the RNA to the 5′ RNA end. The RNA/cDNA duplex fragment mixture, which also includes single-stranded RNA species, is treated to remove primer and the dNTPs, for example, by one or more phenol or ammonium acetate precipitation steps and gel filtration on Sephacryl S-400 (Pharmacia).

The resulting cDNA fragments, either before or after digestion of the RNA strand, are reacted with terminal deoxynucleotide transferase (TdT) in the presence of a selected dNTP, to form a homopolymer region at the 3′ cDNA ends. In the embodiment illustrated in Figure 1, the homopolymer is poly dG, and the homopolymer tailing reaction is performed with RNA/cDNA duplex fragments. Methods for homopolymer tailing have been reported (Jackson; Lobban; Nelson). If the RNA strand is present, the efficiency of the reaction can be enhanced by initial digestion with a 5′ exonuclease, to expose the 3′ fragment ends, or by carrying out the reaction in the presence of cobalt ions (Roychoudhury).

The reaction is carried out under conditions which place at least about 10 and preferably 15 or more bases at the 3′ fragments ends. The reaction described in Example 1, which yields poly dG tailing, is exemplary. If

the tailing reaction is performed with RNA/cDNA fragments, the duplex fragments are now digested with RNase to remove the sense-strand RNA, and the single-strand RNA species in the mixture, yielding single-strand, anti-sense strand cDNAs containing a 5'-end region derived from the specific-sequence primer, the region R of known sequence, and a 3'-end homopolymer tail, e.g., a poly dG tail. The RNase, TdT and selected dNTP are removed, for example, by cDNA fragment precipitation.

Since homopolymer tailing by TdT is effective with both single- and double-stranded material, the RNA strands in the RNA/cDNA duplex fragments can be removed with RNase digestion prior to homopolymer tailing. One advantage of homopolymer tailing applied to the mRNA/cDNA fragments is that the tailing method selects for full-copy cDNA fragments. This is because the terminal transferase reaction with duplex fragments requires that the 3'-end of the cDNA be substantially coterminous with or longer than the 5' end of the RNA strand. Thus a partially copied cDNA will underhang the RNA strand and not be tailed efficiently. As a result, the strand will not be amplified.

For selective amplification, the specific-sequence anti-sense cDNA strands from above are mixed with a homopolymer primer which is homologous to the 3'-end homopolymer tail of the cDNA strands. As used herein, the term "homologous" means that the primer is effective to prime by base-pair specific attachment a primer-binding sequence of a DNA strand, e.g., the 3' homopolymeric region, under the annealing conditions employed. Typically, the primer sequence is the exact complement of the primer binding region, although variations in sequence correspondence may be allowed, particularly for longer primer sequences. In the embodiment illustrated in Figure 1, where the homopolymer is poly dG, the primer includes a poly dC region, containing preferably a run of at least

about 15 dC bases. As will be discussed below, the primer may also contain the sequence of a selected restriction endonuclease site, such as an XhoI site which can be used for fragment cloning. It is noted that homopolymer regions generally have more homopolymer subunits than actually shown in the figures.

It will be appreciated that initial replication of the anti-sense strands, using the homopolymer primer, yields complementary sense strands whose 3' end regions contain a sequence which is complementary to the common sequence provided by the specific-sequence primer. For example, where the specific-sequence primer contains a poly dC sequence, as indicated in Figure 1, the sense strand will terminate with a 3'-end poly dG sequence. The amplification mixture further includes a common-sequence primer which is homologous to the complement of the common-sequence at the 5'-end region of the anti-sense strands. It is noted that the common-sequence primer may be homologous to the known-sequence region of the initial fragment used to form the anti-sense strands, since the sense strands formed from the anti-sense strands will contain this known primer sequence at their 3' ends.

In the method illustrated in Figure 1, the specific-sequence primer includes a 5'-end poly dC sequence which forms the 5' end region of the first-strand (anti-sense) cDNA. The homopolymer which is formed at the 3' end of the anti-sense strand is the complement of the 5'-end sequence, namely poly dG. As will be seen below, the complementary homopolymer ends allow the anti-sense strands to be amplified with a single homopolymer primer which is homologous to the 3'-end strand sequence. That is, the homopolymer primer also serves as the common-sequence primer. One preferred homopolymer primer has the following sequence, which includes a 5'-end XhoI site and two protective G bases:

$$d(5'\text{-GG}\underline{\text{CTCGAG}}\text{C}_{20}\text{-}3')$$
$$\underline{\text{XhoI}}$$

The primer is present in large molar excess with respect to the fragment concentration, and comparable or greater in concentration to the calculated intramolecular strand end concentration, to minimize hybridization between the complementary fragment ends, which can lead to either fragment circles or concatemers. Typically, the primer concentration is at least about 10^3 fold higher than cDNA fragment concentration, and about 1uM or greater. It may also be necessary to supplement the reaction with added primer during amplification, to replenish primer used up in strand replication.

Purified anti-sense cDNA strands from above are amplified by repeated fragment duplication according to the following steps. The fragments are mixed with a large molar excess of the homopolymer primer described above. As noted above, the primer sequences and lengths are such as to allow primer attachment to the 3'-end homopolymeric regions of the sense and anti-sense strands, sufficient for DNA polymerase priming. Another requirement is that the primer, when attached to the linker region of the denatured fragment strands, be capable of priming polymerase-catalyzed strand replication; that is, that the internal end of the primer provides a free 3'-OH.

The denatured fragments and primer are initially cooled to a temperature, typically between about 37-60°C, to allow homopolymer primer attachment to the anti-sense cDNA fragments. The four dNTPs and a DNA polymerase capable of catalyzing second-strand primed replication are added, and the reaction mixture is brought to a temperature suitable for enzymatic strand replication.

In the method described in Example 1 below, the dNTPs and polymerase were added prior to fragment denaturation. After heat denaturing, the mixture was annealed at 50°C for two minutes, then brought to 72°C for 5-12 minutes for primed, second-strand replication. The DNA polymerase used was <u>Thermus aquaticus</u> DNA polymerase (Taq DNA polymerase), which is relatively heat-stable at up to 95°C for brief periods.

It will be appreciated from the above, and from Figure 1, that the homopolymer primer hybridizes with the 3'-end homopolymer regions in each DNA strand, and thus a doubling of fragment number occurs with each replication. The above replication procedure, which involves fragment denaturation, annealing to form fragment strand primer complexes, and second strand replication of the complexes in the presence of DNA polymerase, is repeated until a desired concentration of fragments is achieved. In the above example, which employs a heat-stable <u>Taq</u> polymerase, the three replication steps are carried out simply by heating the fragment mixture to a denaturing temperature (above the T_m of the fragments), cooling briefly to allow fragment/primer complex formation, and incubating for a period sufficient for second-strand synthesis. The period of polymerization may be extended for up to about one-half hour for longer fragments, e.g., up to about 8 kilobases in length.

Since the concentration of fragments doubles at each round of replication, a 10^3 fold amplification can be achieved with about 10 rounds of replication, and a 10^6 fold amplification with about 20 rounds of replication. Thus, initial picogram amounts of sequence-specific RNA species will yield microgram amounts of fragment material after about 20 rounds of replication.

Example 1 illustrates the application of the amplification method to a mixture of poly A RNA species

isolated from human peripheral T cells. The specific-primer sequence used for initial cDNA synthesis is a human interleukin-2 gene sequence having a 5'-end poly dC sequence.

Following the amplification procedure, the selectively amplified fragments are separated from the polymerase and polynucleotide components, typically by phenol extraction. If the homopolymer primer used for amplification includes a 5'-end restriction site sequence, the fragments can be cut with the corresponding endonuclease, and cloned into a compatible insertion site in a suitable vector, according to known cloning procedures.

Figure 2 illustrates another embodiment of the invention, in which selective amplification results in duplex DNA fragments having different selected restriction sites at opposite fragment ends, to allow directional cloning of the amplified fragments.

The method differs from the single-primer method above in that the opposite ends of the anti-sense strands which are amplified have non-complementary sequences. As will be seen, amplification requires two different-sequence primers, i.e., the homopolymer and common-sequence primers have different sequences. These two different primers may be further distinguished by different 5'-end restriction site sequences, to produce different selected restriction sites at the opposite ends of the amplified fragments.

In the specific method illustrated in Figure 2, the specific-sequence primer used for first-strand cDNA synthesis contains a 5' poly dA sequence which will become the 5'-end common sequence of the anti-sense strand cDNAs. After first-strand synthesis, the fragments are treated with TdT in the presence of a selected deoxynucleoside triphosphate, e.g., dGTP, to form a homopolymer tail,

e.g., poly dG, at the first strand cDNA 3′ end. The only requirement in the selection of the 3′-end homopolymer is that it not be complementary to the 5′-end common sequence. Preferably, one end of the anti-sense fragment has a poly dA or poly dT sequence, and the opposite end, a poly dC or poly dG sequence.

After digestion of the RNA strand and removal of the reverse transcription reaction components, the anti-sense strands are mixed with a homopolymer primer which includes a 5′-end sequence corresponding to one selected restriction site, and a common-sequence primer which includes a 5′-end sequence corresponding to a second selected restriction site, where the primer restriction sequences preferably correspond to sites not present in the known sequence region of the fragment being amplified. The two primers may further include protective bases at their 5′ ends. An exemplary homopolymer primer having a 5′ XbaI site and a common-sequence primer having a 5′ XhoI site are shown at A and B below, respectively:

(A) $d(5'\text{-GG}\underbrace{\text{TCTAGA}}_{\textit{XbaI}}\text{C}_{20}\text{-}3')$

(B) $d(5'\text{-GG}\underbrace{\text{CTCGAG}}_{\textit{XhoI}}\text{A}_{20}\text{-}3')$

The anti-sense fragments, which represent only those original RNA species having the known selected sequence, are amplified as above, with repeated cycles of primer annealing, polymerization, and denaturation. At each cycle, the homopolymer primer is used to prime the anti-sense strand, and the common-sequence primer, the sense strand.

For cloning, the fragments are digested with the two restriction endonucleases formed at the amplified fragment ends, and ligated into a suitable cloning vector

which has directional ends which are compatible with the fragment ends. The cloning method described in Example 1 is illustrative. Here the fragments are formed with rare fragment-end restriction sites XbaI and XhoI, digested with the two restriction enzymes and ligated into a Bluescript™ vector having unique XbaI and XhoI sites. Ligation and selection of successful recombinants is conventional.

Preferred vectors for directional cloning contain an RNA polymerase promoter site adjacent the insertion site which corresponds to the 5′ end of the amplified RNA. When this vector is cut at the restriction site at the fragment insert end opposite the promoter, and treated with RNA polymerase in the presence of all four NTPs, the system produces RNA strands which can be used by the host for protein production. It will be appreciated that the two primer restriction sites must be selected such that insertion into the vector places the 5′ end of the anti-sense strand in the duplex cDNA adjacent the polymerase promoter. Alternatively, the cloning vector may be provided with RNA polymerase promoters at both insertion sites, allowing for sense-strand transcription by cutting the recombinant vector at the fragment end which insures sense-strand transcription.

II. Selective Amplification: DNA Fragments

The invention is also useful for selective amplification of DNA fragments, typically genomic DNA fragments, containing a region of known sequence. Genomic DNA from a selected cell source can be isolated by standard procedures, which typically include successive phenol and phenol/chloroform extractions with ethanol precipitation. The isolated DNA may be obtained from isolated chromosomes or chromosomal regions of interest. The duplex DNA is fragmented preferably by partial or

complete digestion with one or more selected restriction endonucleases, although mechanical shearing may be employed. The fragmented genomic pieces may be size fractionated, or further treated to remove repetitive DNA. Other sources of single- or double-stranded DNA fragments can include extrachromosomal material, e.g., mitochondrial DNA, double-stranded DNA viruses, or viruses which have as part of their life cycle a double-stranded intermediate, e.g., retroviruses.

Linearized or fragmented plasmid DNA, or fragmented phage DNA is another source of DNA fragments which one might wish to amplify. The vector DNA is obtained from purified plasmid or phage DNA according to conventional techniques, and linearized and/or fragmented by digestion with selected restriction endonuclease(s).

The selective amplification method, which is illustrated in Figure 3, generally follows the amplification method described above, except that the initial sense-strand template used to form the sequence-specific anti-sense strands is single-stranded DNA rather than RNA.

In practicing the method, the duplex fragments are first treated to block the 3'-end OH groups of the duplex fragments. This insures that homopolymer tailing occurs only at the 3'-ends of newly synthesized (with sequence-specific priming) DNA strands. Blocking of the 3' fragment ends may be carried out using T4 DNA polymerase and cordycepin triphosphate (3'-deoxy ATP) using conditions similar to those described in Maniatis (pp. 118-119) for repair of 3' DNA ends.

The initial duplex fragment mixture includes at least a small portion of DNA fragments containing the region R with a known sequence. The duplex fragments are denatured, preferably by heating, and mixed with a sequence-specific primer such as described above containing a 3' region (shown as a sawtooth line in the figure)

which is homologous to the known sequence region. Where, as in Figure 3, the amplification uses same-sequence primers, the sequence-specific primer also includes a homopolymeric 5' sequence which is homologous to the homopolymer tail to be formed at the 3' ends of the newly made anti-sense strands. The primed fragments are copied with DNA polymerase and all four dNTPs, producing a new DNA strand having the common-specific primer sequence at its 5' end. This new strand is also referred to herein as an anti-sense strand, in analogy to first-strand cDNA synthesis.

After removing the polymerase and four dNTPs, the fragments are treated with TdT and a selected dNTP, to produce homopolymer tailing at newly synthesized strands in the duplex fragments, as indicated in the figure. In the example shown, where the first-strand primer includes a homopolymer binding region, such as poly dC, the homopolymer tail is preferably complementary to the primer sequence, such as poly dG, allowing amplification with same-sequence primers. Selective amplification of the known-sequence fragments is carried out as above, by repeated cycles of primer annealing, polymerization, and denaturation, until a desired degree of amplification is achieved. The method is illustrated in Example 2.

It will be appreciated from Figure 2 how the fragment amplification method using different-sequence primers can be applied to duplex DNA fragments. In particular, the invention contemplates using different-sequence homopolymer and common-sequence primers for producing amplified fragments with directional restriction sites, as in the method illustrated in Figure 2.

III. Utility

One application of the selective amplification method of the invention is for identifying and/or isolat-

ing from a fragment mixture, RNA or DNA species having a known sequence. The method is useful as a diagnostic tool for identifying the presence of an infective or other pathogenic agent in a fragment mixture obtained from a cellular, tissue, or body-fluid source. In this application, the nucleic acid fragments are extracted from the source, and those fragments with the selected known sequence are amplified as above. The amplified fragments can be detected by conventional means, such as hybridization with sequence-specific probes and/or can be isolated and purified, such as by hybridization with sequence-specific biotinylated probes, coupled with binding to an affinity solid support.

The method may also be used for cloning specific-sequence RNA or DNA species, particularly where these are present in low amounts and/or represent only a very small portion of the total fragment material. The method exploits the ability to place selected restriction sites at the ends of amplified fragments for directional or non-directional cloning. Thus, for example, the method may be used to amplify the coding sequence of a selected mRNA species, from a RNA fragment mixture, and to insert the fragment into a suitable expression vector.

Another application of the invention, for identification and analysis of restriction fragment length polymorphisms (RFLPs), is illustrated in Figure 4. The two DNA fragments at the top in the figure represent genomic DNA fragments which each have a region R of common sequence, but are distinguished by a mutation which has resulted in a new restriction site, indicated by RE(new), which has the same sequence as existing restriction site RE, but which is significantly closer to the region of known sequence, as indicated. This restriction site polymorphism may, for example, be diagnostic of a given

genetic disease, or may be useful as a marker for determining predisposition to a given genetic disease.

The fragments are first digested to completion with the restriction endonuclease which cuts at the RE and RE(new) sites, thus producing fragments which terminate at the sites of interest. The fragments are now treated to block 3'-OH groups, as above, denatured and polymerized in the presence of a specific-sequence primer which can bind to the region R sequence and which contains a common sequence. As indicated in the figure, the resulting newly synthesized strands extend from a 5'-end primer sequence to the restriction site of interest at the 3' strand ends. Homopolymer tailing now selectively adds a homopolymer sequence to the 3'-ends of the newly synthesized strands. The fragments are amplified, as above, in the presence of same-sequence or different-sequence primers. In the method illustrated in Figure 4, the strand ends are formed, as above, for same-sequence primer amplification.

The sizes of the amplified fragments are now examined, typically by size fractionation on agarose gel and Southern blotting using a radiolabeled probe. The results of an idealized gel pattern are shown at the bottom in the figure. From this gel pattern, it can be determined that a restriction fragment size polymorphism exists, and the sizes of the two fragments can be measured.

The following examples illustrate the method of fragment amplification and fragment isolation described above, but are in no way intended to limit the scope of the method or its applications.

Materials

Bluescript™ is obtained from Stratagene (San Diego, CA). Terminal deoxynucleotide transferase (calf thymus), alkaline phosphatase (calf intestine),

polynucleotide kinase, E. coli DNA polymerase I (Klenow fragment), and S1 nuclease are obtained from Boehringer Mannheim Biochemicals (Indianapolis, IN);

XbaI, XhoI, T4 DNA ligase and T4 DNA polymerase are obtained from New England Biolabs (Beverly, MA); and streptavidin agarose, from Bethesda Research Labs (Bethesda, MD). Low-gelling temperature agarose (Sea Plaque) is obtained from FMC (Rockland, ME). Nitrocellulose filters are obtained from Schleicher and Schuell (Natick, MA).

Synthetic oligonucleotide linkers and primers are prepared using commercially available automated oligonucleotide synthesizers. Alternatively, custom designed synthetic oligonucleotides may be purchased, for example, from Synthetic Genetics (San Diego, CA). The cDNA synthesis kit and random priming labeling kits were obtained from Boehringer-Mannheim Biochemical (BMB, Indianapolis, IN).

Site specific DNA cleavage is performed by treating with the suitable restriction enzyme (or enzymes) under conditions which are generally understood in the art, and the particulars of which are specified by the manufacturer of these commercially available restriction enzymes. See, e.g., New England Biolabs, Product Catalog. In general, about 1 ug of plasmid or DNA sequence is cleaved by one unit of enzyme in about 20 ul of buffer solution after 1 hr digestion at 37°C; in the examples herein, typically, an excess of restriction enzyme is used to insure complete digestion of the DNA substrate.

Incubation times of about one hour to two hours at about 37°C are workable, although variations can be easily tolerated. After each incubation, protein is removed by extraction with phenol/chloroform, and may be followed by ether extraction, and the nucleic acid recovered from aqueous fractions by precipitation with

ethanol (70%). If desired, size separation of the cleaved fragments may be performed by polyacrylamide gel or agarose gel electrophoresis using standard techniques. A general description of size separations is found in Methods in Enzymology (1980) 65:499-560.

Restriction cleaved fragments may be blunt ended by treating with the large fragment of *E. coli* DNA polymerase I (Klenow reagent) in the presence of the four deoxynucleoside triphosphates (dNTPs) using incubation times of about 15 to 25 min at 20° to $25^{\circ}C$ in 50 mM Tris pH 7.6, 50 mM NaCl, 6 mM $MgCl_2$, 6 mM DTT and 0.1-1.0 mM dNTPs. The Klenow fragment fills in at 5' single-stranded overhangs in the presence of the four nucleotides. If desired, selective repair can be performed by supplying only one of the, or selected, dNTPs within the limitations dictated by the nature of the overhang. After treatment with Klenow reagent, the mixture is extracted with phenol/chloroform and ethanol precipitated. Treatment under appropriate conditions with S1 nuclease results in hydrolysis of any single-stranded portions of DNA. In particular, the nicking of 5' hairpins formed on synthesis of cDNA is achieved.

Blocking of 3' ends of DNA is carried out using T4 DNA polymerase and cordycepin triphosphate (3'-deoxy ATP) using conditions similar to those described by Maniatis (p. 118-119) for repair of 3' ends of DNA. Only cordycepin is used for repair.

Ligations are performed in 15-50 ul volumes under the following standard conditions and temperatures: for example, 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 33 mg/ml BSA, 10 mM-50 mM NaCl, and either 40 mM ATP, 0.01-0.02 (Weiss) units T4 DNA ligase at $14^{\circ}C$ (for "sticky-end" ligation) or 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at $14^{\circ}C$ (for "blunt-end" ligation). Intermolecular "sticky-end" ligations are usually

performed at 33-100 mg/ml total DNA concentrations (5-100 nM total end concentration). Intermolecular blunt-end ligations are performed at 1 mM total ends concentration.

In vector constructions employing "vector fragments", the vector fragment is commonly treated with bacterial alkaline phosphatase (BAP) or calf intestinal alkaline phosphatase (CIP) in order to remove the 5′ phosphate and prevent self-ligation of the vector. Digestions are conducted at pH 8 in approximately 10 mM Tris-HCl, 1 mM EDTA using about 1 unit per mg of BAP at 60°C for one hour or 1 unit of CIP per mg of vector at 37°C for about one hour. In order to recover the nucleic acid fragments, the preparation is extracted with phenol/chloroform and ethanol precipitated. Alternatively, religation can be prevented in vectors which have been double digested by additional restriction enzyme digestion and separation of the unwanted fragments.

Example 1
Selected Sequence mRNA Amplification

A. cDNA Synthesis

Human peripheral blood T cells are isolated from normal individuals by T rosetting, and total RNA is isolated from about 10^7 cells according to standard procedures (Cathula). The total RNA preparation is fractionated by oligo dT chromatography, also according to known procedures (Maniatis, p. 211), yielding a poly A mRNA preparation. The final preparation is diluted to an RNA concentration of about 1 ug/ml.

The mRNA fragments are mixed with about 1 uM final concentration of a primer containing a 5′-region 20-mer sequence of the human interleukin-2 gene corresponding to the coding region bases 381-400 of exon 1 (Holbrook), and an 18-mer 3′-region poly dC sequence. The primer is

prepared by standard oligonucleotide synthesis and has the sequence:

d(5'-C_{18}ATGCTCCAGTTGTAGCTGTG-3')

The fragment/primer mixture is converted to first-strand cDNA, using conditions described in the cDNA kit supplied by Boehringer-Mannheim.

The RNA/cDNA complex is purified by sequential extraction with one volume of phenol/chloroform followed by 2 ethanol (5 ug of tRNA is added as carrier) precipitations using 0.5 volume of 7.5 M NH_4OAc pH 7.0 and 2 volumes of ethanol. After centrifugation (10 min in microfuge at room temperature; about 14K rpm) the pellet is completely desiccated in a speed vac to assure complete removal of the NH_4OAc which is known to inhibit TdT.

The fragment mixture is resuspended in 0.050 ml of 0.2 M potassium cacodylate (pH 7.2), 4 mM $MgCl_2$, 1 mM 2-mercaptoethanol, as provided by BRL, which contains a final concentration of 25 uM of dGTP. To initiate the reaction, 15 units of TdT (supplied by BRL) is added and the reaction is incubated for 30 min at 37°C. At the end of the reaction 10-20 ug of RNase A (supplied by Boehringer-Mannheim) is added to the reaction and incubated at 37°C for an additional 30 min. At the conclusion of the RNase A reaction, the reaction mix is extracted with phenol/chloroform as above and 5-10 ug of tRNA is added as carrier before ethanol precipitation, as described above was carried out. Alternatively, the sample may be precipitated without carrier and treated with RNase A after resuspending the samples.

B. Amplifying the cDNA fragments

To 100 ul of 10 mM Tris-Cl buffer, pH 8.3, containing 1.5 mM $MgCl_2$ (Buffer A) is added 100 ul of the

above cDNA fragment mixture 2 uM of a primer having the sequence d(5'-GGCTCGAGC$_{20}$-3'), 200 uM (final concentration) each of dATP, dCTP, dGTP, and dTTP, and 5 units of Thermus aquaticus DNA polymerase (Taq polymerase). The reaction mixture is heated to 94°C for 1 minute for denaturation, allowed to cool to 50°C for 2 min for primer annealing, and then heated to 72°C for 5-12 min to allow for primer extension by Taq polymerase. The replication reaction, involving successive heating, cooling, and polymerase reaction, was repeated an additional 25 times with the aid of a Perkin Elmer Cetus DNA thermal cycler.

C. Cloning Amplified Fragments

The amplified fragments from the sample mixture corresponding to 10^5 cells are digested with XhoI. Bluescript™ M13 plasmid is treated with the same enzyme, and treated with alkaline phosphatase, prior to mixing with the above amplified fragments. Ligation is performed under conditions which promote circularization of the single plasmid fragments. The circularized plasmid is selected on E. coli strain DH5, and successful recombinants are selected for ampicillin resistance.

While the invention has been described with reference to particular embodiments, methods, construction and use, it will be apparent to those skilled in the art that various changes and modifications can be made without departing from the invention.

IT IS CLAIMED:

1. A method of selectively amplifying nucleic acid fragments having a region of known sequence, comprising

combining fragment sense strands containing the known sequence with a sequence-specific primer having a 3' sequence which is homologous to the known sequence, under conditions in which the primer binds to the sense strands,

reacting the primed sense strands with DNA polymerase in the presence of all four deoxynucleotide triphosphates, under conditions which produce copying of the sense strands, to form anti-sense strands having a common 5'-end primer sequence,

treating the anti-sense strands with terminal deoxynucleotide transferase and a selected deoxynucleoside triphosphate, to add a homopolymeric sequence to the 3' ends of the anti-sense strands,

mixing the anti-sense fragments with a homopolymer primer which is homologous to said homopolymeric fragment sequence, a common-sequence primer which is homologous to the complement of the 5'-end common primer sequence, DNA polymerase and all four deoxynucleoside triphosphates,

annealing the mixture under conditions which allow fragment priming,

reacting the mixture under conditions in which the primed fragments are converted to double-stranded fragments,

denaturing the fragments, and

repeating said, annealing, reacting, and denaturing steps until a desired degree of fragment amplification has been achieved.

2. The method of claim 1, for amplifying duplex DNA fragments having a region of known sequence, which further includes blocking the 3'-end OH groups of the fragments prior to said treating, to prevent homopolymer tail addition to either strand of the duplex fragments.

3. The method of claim 1, wherein the specific-sequence primer contains a 5' sequence which is complementary to the homopolymeric sequence added at the 3' end of the anti-sense strand, and said homopolymer primer also serves as the common-sequence primer.

4. The method of claim 1, for use in identifying restriction fragment length polymorphisms in duplex DNA fragments, each having a region with the same known sequence, and specific restriction endonuclease sites which are spaced at different distances from said known sequence region, which further includes digesting the duplex DNA fragments with the corresponding restriction endonuclease, size fractionating the fragments after said repeating, and identifying the size characteristics of those fragments which have been selectively amplified.

5. The method of claim 1, for use in determining the distance between the known sequence of a duplex DNA fragment and an adjacent restriction endonuclease site, which further includes digesting the duplex DNA fragment with the corresponding restriction endonuclease, size fractionating the fragment after said repeating, and determining the size of the amplified fragment.

6. The method of claim 1, for use in producing selectively amplified duplex fragments with a selected restriction endonuclease site at the fragment ends,

wherein the homopolymer primer and the common-sequence primer each includes a 5'-end restriction endonuclease sequence, and the amplified fragments and cloning vector are both digested with the corresponding endonuclease, mixed, and ligated under fragment-insertion conditions.

7. The method of claim 1, for use in producing selectively amplified cDNA fragments having first and second different selected restriction endonuclease sites at opposite fragment ends, wherein the homopolymer primer includes a 5'-end sequence corresponding to the sequence of the first selected restriction site, 5' with respect to the binding sequence, and the common-sequence primer includes a 5'-end sequence corresponding to the second selected restriction site.

8. The method of claim 1, for use in selectively amplifying anti-sense strand fragments derived from an RNA having a region of known sequence, wherein the sense strand fragments are RNA species, said initial reacting produces RNA/cDNA duplex fragments, and which further includes digesting the RNA/cDNA duplex fragments with RNase to remove the RNA strand.

9. The method of claim 8, wherein the specific sequence primer contains a 5'-end sequence which is complementary to the homopolymeric sequence added at the 3' end of the anti-sense strand, and said homopolymer primer also serves as the common-sequence primer.

AMENDED CLAIMS
[received by the International Bureau
on 28 November 1989 (28.11.89)
new claim 2 added; original claim 2 unchanged but
renumbered as claim 3; claims 3-9 amended and
renumbered as claims 4-10 (3 pages)]

1. A DNA amplification method comprising

preparing a duplex nucleic acid molecule having sense and anti-sense strands, where the 5'-end of the anti-sense strand terminates in a known sequence,

treating the anti-sense strand with terminal deoxynucleotide transferase and a selected deoxynucleoside triphosphate, to add a homopolymeric end sequence to the 3'-end of the anti-sense strand,

mixing the homopolymeric-tailed anti-sense strand, under conditions in which the anti-sense strand is in a single-stranded form, with a homopolymer primer which is complementary to said homopolymer end sequence, a known-sequence primer which is homologous to the complement of the known terminal 5'-end sequence of the anti-sense strand, DNA polymerase, and all four deoxyribonucleoside triphosphates,

annealing the mixture under conditions which allow priming of the anti-sense strand,

reacting the mixture under conditions in which the primed fragments are converted to double-stranded fragments,

denaturing the fragments, and

repeating said annealing, reacting, and denaturing steps until a desired degree of amplification of fragments containing the anti-sense strand has been achieved.

2. The method of claim 1, wherein the 3'-end of the sense-strand of the duplex nucleic acid which is prepared is unable to act as a substrate for terminal deoxynucleotide transferase.

3. The method of claim 2, wherein the duplex nucleic acid molecule which is prepared is a duplex DNA in which the sense strand is blocked at its 3'-end hydroxyl group.

4. The method of claim 1, wherein the known 5'-end sequence of the anti-sense strand contains a terminal 5' sequence which is complementary to the homopolymeric sequence added at the 3'-end

of the anti-sense strand, and said homopolymer primer also serves as the known-sequence primer.

5. The method of claim 1, for use in identifying restriction fragment length polymorphisms in duplex DNA fragments, each having a region with the same known sequence, and specific restriction endonuclease sites which are spaced at different distances from said known sequence region, wherein said preparing includes digesting the duplex DNA fragments with the corresponding restriction endonuclease, and which further includes size fractionating the amplified fragments and identifying the size characteristics of those fragments which have been selectively amplified.

6. The method of claim 1, for use in determining the distance between the known sequence of a duplex DNA fragment and an adjacent restriction endonuclease site, wherein said preparing includes digesting the duplex DNA fragment with the corresponding restriction endonuclease, and which further includes size fractionating the amplified fragments and determining the size of the amplified fragment.

7. The method of claim 1, for use in producing selectively amplified duplex fragments with a selected restriction endonuclease site at the fragment ends, wherein the homopolymer primer and the known-sequence primer each includes at its 5'-end, a restriction endonuclease recognition sequence, and the amplified fragments are digested with the corresponding endonuclease.

8. The method of claim 1, for use in producing selectively amplified cDNA fragments having first and second different selected restriction endonuclease sites at opposite fragment ends, wherein the homopolymer primer includes a 5'-end sequence corresponding to the sequence of the first selected restriction site, 5' with respect to the homopolymer binding sequence, and

the known-sequence primer includes a 5'-end sequence corresponding to the second selected restriction site.

9. The method of claim 1, for use in selectively amplifying anti-sense strand fragments derived from single-strand RNA having a region of known sequence, wherein said preparing includes using said RNA as a template to form an RNA/cDNA duplex nucleic acid molecule.

10. The method of claim 9, wherein the known terminal 5'-end sequence of the anti-sense strand contains a 5'-end homopolymer sequence which is complementary to the homopolymeric sequence added at the 3'-end of the anti-sense strand, and said homopolymer primer also serves as the known-sequence primer.

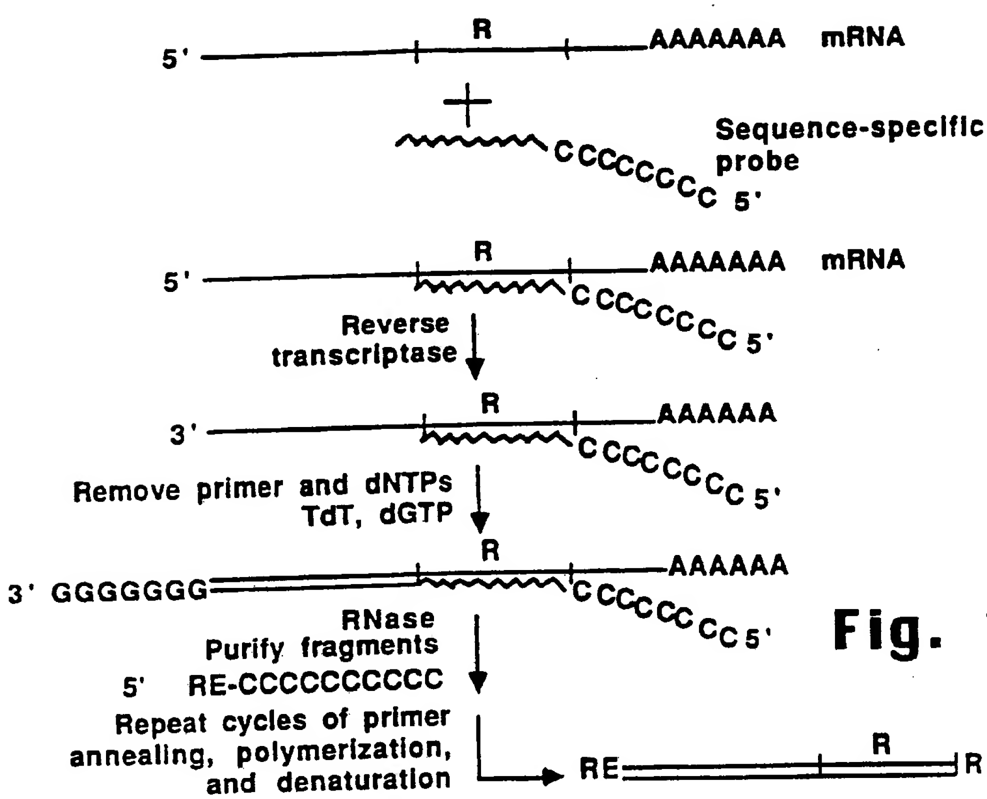
5'
R
AAAAAAA mRNA
Sequence-specific probe
CCCCCCCC 5'
5'
R
AAAAAAA mRNA
CCCCCCCC 5'
Reverse transcriptase
3'
R
AAAAAA
CCCCCCCC 5'
Remove primer and dNTPs
TdT, dGTP
3' GGGGGGG
R
AAAAAA
CCCCCCCC 5'
RNase
Purify fragments
5' RE-CCCCCCCCCC
Repeat cycles of primer annealing, polymerization, and denaturation
RE
R
RE

Fig. 1

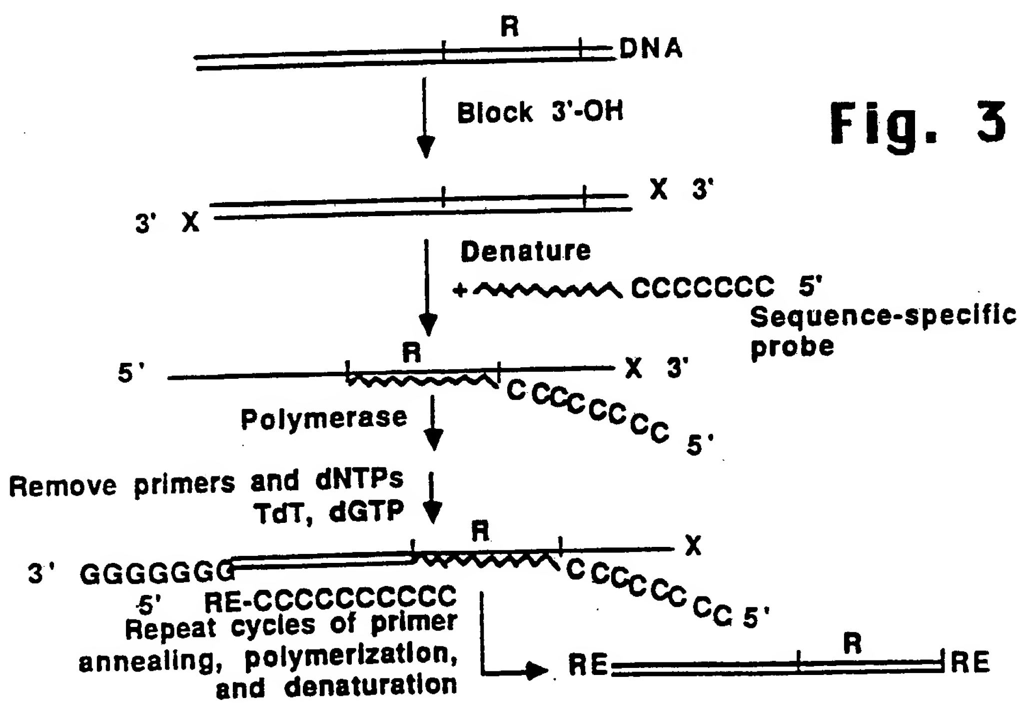
R
DNA
Block 3'-OH
3' X
X 3'
Denature
+ CCCCCCC 5'
Sequence-specific probe
5'
R
X 3'
CCCCCCCC 5'
Polymerase
Remove primers and dNTPs
TdT, dGTP
3' GGGGGGG
R
X
CCCCCCCC 5'
5' RE-CCCCCCCCCC
Repeat cycles of primer annealing, polymerization, and denaturation
RE
R
RE

Fig. 3

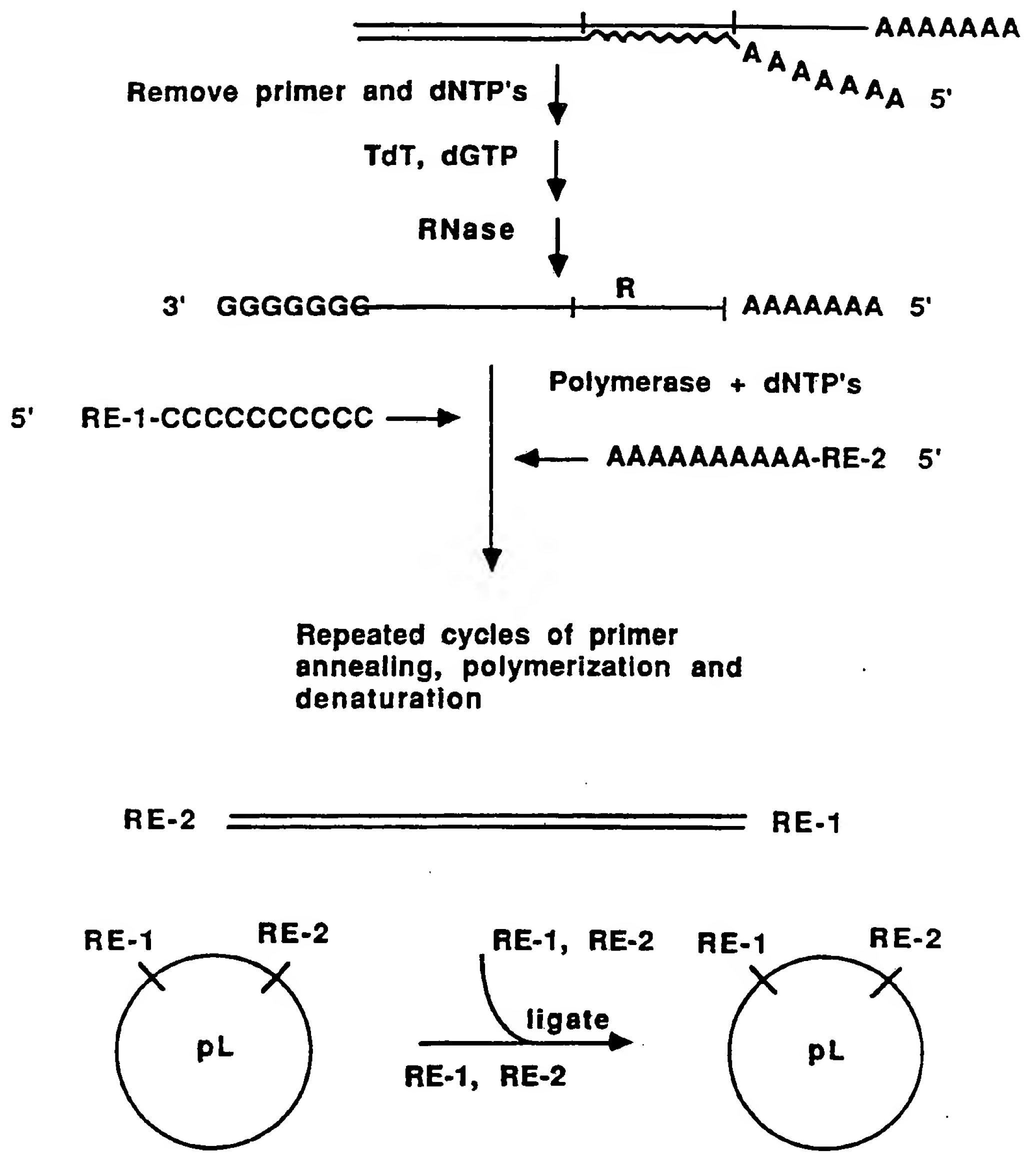
AAAAAAA
AAAAAAA 5'
Remove primer and dNTP's
TdT, dGTP
RNase
R
3' GGGGGGG AAAAAAA 5'
Polymerase + dNTP's
5' RE-1-CCCCCCCCCC
AAAAAAAAAA-RE-2 5'
Repeated cycles of primer annealing, polymerization and denaturation
RE-2 RE-1
RE-1 RE-2
pL
RE-1, RE-2
ligate
RE-1, RE-2
RE-1 RE-2
pL

Fig. 2

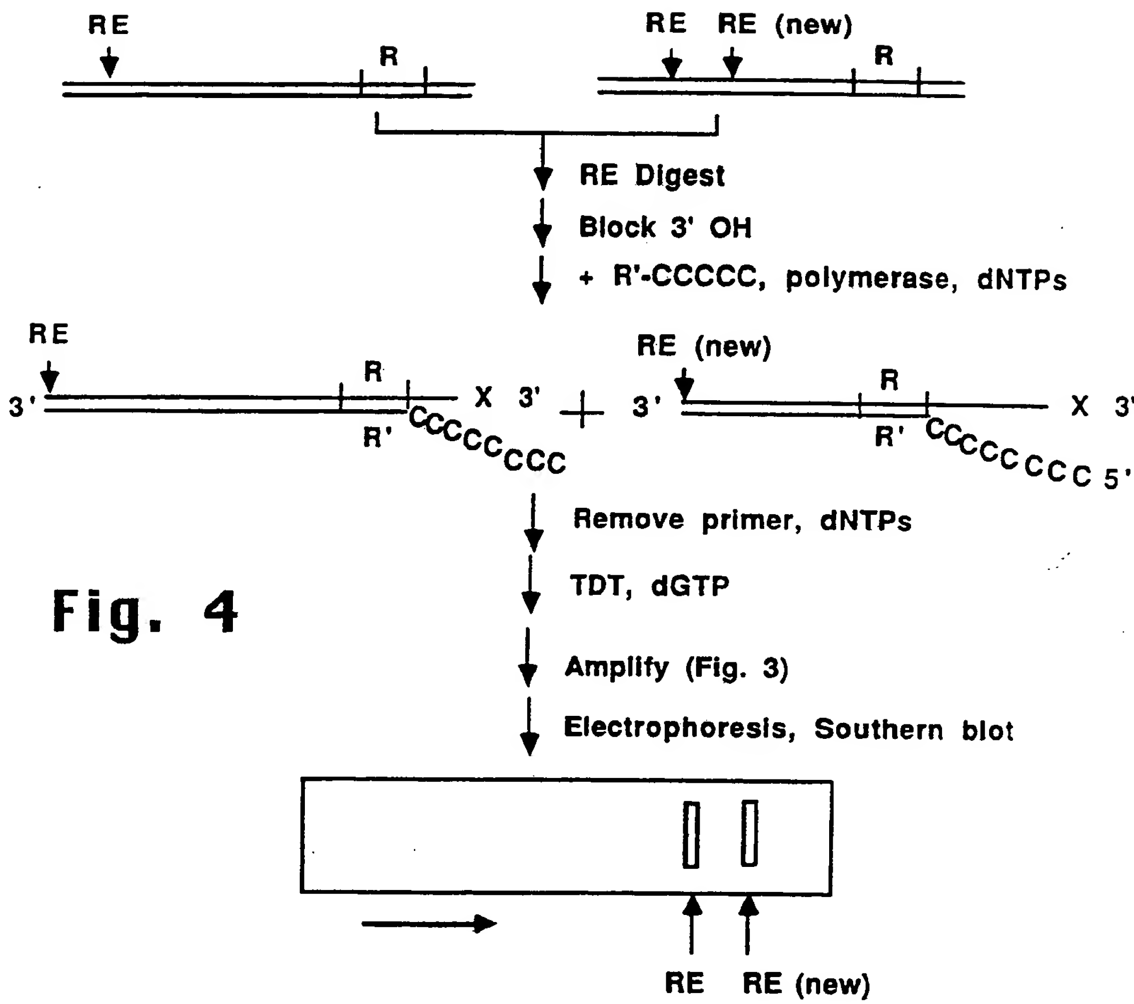
RE
R
RE
RE (new)
R
RE Digest
Block 3' OH
+ R'-CCCCC, polymerase, dNTPs
RE
R
3'
X 3'
R'
CCCCCCCCC
RE (new)
R
3'
X 3'
R'
CCCCCCCCC 5'
Remove primer, dNTPs
TDT, dGTP
Amplify (Fig. 3)
Electrophoresis, Southern blot
RE
RE (new)

Fig. 4

INTERNATIONAL SEARCH REPORT

International Application No PCT/US89/03099

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

INT. CL(4): C12P 19/34; C12N 15/00; C12Q 1/68
U.S. CL: 435/6, 91, 172.3

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.	435/6,91,172.3 935/17,18,76,77,78

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

APS, JPOABS: Nucleic acid amplif?
DIALOG(MEDICINE, BIOTECH): Polymerase chain reaction, restriction fragment length polymorphisms

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No [13]
Y	US,A 4,683,202 MULLIS 28 July 1987	
	Columns 2-3	1
	Column 5 and 8	8
	Column 6	3,9
	Diagrams, columns 9-12	1,3,7,9
	Column 13	4
	Column 14	7
Y	US,A 4,683,195 MULLIS 28 July 1987	
	Columns 3 and 9	1
	Columns 7 and 13	8
	Column 8	1,7,6
	Diagrams, columns 11-14	1,3,7,9
	Column 15 and 18	6
	Column 18 and 38	4

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
15 August 1989	20 SEP 1989
International Searching Authority	**Signature of Authorized Officer**
ISA/US	Stephanie W. Zitomer Dr. Stephanie W. Zitomer

International Application No PCT/US89/03099

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
Y,P	US,A 4,800,159 MULLIS 24 January 1989	
	Columns 2 and 8	1
	Column 7	1,7,8
	Column 10	8
	Diagrams, Columns 11-14	1,3,7,9
	Columns 14-15 and 17	6
	Column 18	4
Y	US,A 4,683,194 SAIKI 28 July 1987	
	Column 3	4-7
	Column 6	4-8
	Column 7	4,5
Y,P	Science Vol. 243, 1989. LOH "Polymerase chain reaction with single-sided specificity: analysis of T cell receptor sigma chain." pages 217-218.	1,3,4, 8,9
Y	Molecular Cloning: A Laboratory Manual MANIATIS Cold Spring Harbor Lab., 1982. pages 213, 217, 222 and 374-375	8,1,6, 4-5